Part II

Competing Approaches
Terminology (Schimtz, 2006)

1. Terminology: Set of designations belonging to one special language.

2. Discipline that studies the structure, formation, development, usage and management of terminologies in various subject fields.
2nd Approach Terminology and Information Tools

1. Terminology (1930s): national projects (1960s)

2. Terminography (1976) (Linguistics)

Terminology: National Projects

The making of terminological dictionaries focuses on presenting an unambiguous relationship between terms, concepts and objects

Standardized and Harmonized Terms
Triadic Relationship (Odgen & Richards, 1923)

- Objects: entities in the real world
- Concepts: units of knowledge
- Terms: verbal designations of concepts
Practical Work: Records (Schimtz, 2006)

1. Concept-related categories
2. Term-related categories
3. Administrative-related categories
Concept-related categories

- Definition
- Subject field
- Illustration (symbol or formula)
- Classification
- Superordinate concept
- Subordinate concept
- Co-ordinate concept
Term-related categories

- Synonyms
- Abbreviated forms
- Orthographical variants
- Context (example)
- Grammar data
- Variants (geographical, register)
- Project code
Administrative categories

• Identification number
• Date of creation (adaptation)
• Author
• Source
• Reliability
• comments
Examples

InterActive Terminology for Europe

carrying amount
accounting
Problems (1970s)

(i) Synonymy
(ii) Polysemy
(iii) Homonymy
(iv) Variation
(v) Anisomorphism
(vi) Idiosyncrasy (muldidimensional perspective)
Solutions

(i) Terminography (1970s)

(ii) Knowledge Engineering (2000s)
Assumptions

• Linguistics (especially, Corpus Linguistics)
• Knowledge Management
• Computer Science
• Information Science
Examples

1. **Basic Concepts and Best Practices:**
   - Applied *Linguistics*

Virginia Tech Multimedia Music Dictionary? Why is Linguistics needed for compiling this dictionary?
Critical Remarks

Terminology is a “language discipline dedicated to the scientific study of the concepts and terms used in specialized languages.”
Critical Remarks

Terminological work needs:

1. Describing conceptual relations

2. Describing a term’s combinatorial potential (Yes, but Time Cost?)
Critical Remarks

1. Present conceptual Organizations
2. Offer the multidimensional nature of terms
3. Extract semantic and syntactic information through the use of corpora
Critical Remarks

Conceptual Organizations must reflect some kind of *internal structure*, usually:

- Generic-specific;
- Whole-to-part
- Associative relations
Critical Remarks

Multidimensionality: Term description(s) can be approached from several aspects or dimensions (partly agree)
Critical Remarks

• (semi-)automatic extraction of semantic and syntactic information through the use of corpora, especially from the Internet used as a corpus.

• Methods: working in knowledge-rich contexts
Results

• Terminological Products do not show what the theory says they must show

• Results:
  – Glossaries
  – Vocabularies
  – Terminology Databases (Data Bank)
  – Knowledge bases
TERMIUM

Plus®

The Government of Canada’s terminology and linguistic data bank.

TERMIUM Plus®
TERMIUM Plus®

carrying amount (database)

accounting (Translation Bureau)
National Languages

DANTERM
Ordbog

TERMCAT
Cercaterm: accounting
Individual Projects

• Communicative Terminology:
  – Genome

• Frame-based Terminology:
  – OncoTerm & EcoLexicon

• Termontography:
  – FF POIROT
Communicative Terminology

• Terms are multidimensional
• Terms behave as words do
• Terms contain three dimensions:
  – Linguistic dimension
  – Cognitive dimension
  – Communicative dimension
GENOME

Knowledge database on Human Genome:

- Terminological bank
- Textual corpus
- Document bank
- Picture bank
Frame-based Terminology

Frame-based Linguistics:

– Conceptual organization
– Multidimensionality
– Extraction of semantic and syntactic information through the use of multilingual corpora
OncoTerm: Frame Based Terminology

Sistema Bilingüe de Información y Recursos Oncológicos (2002)

1,896 related concepts (meaning/conceptual relations)

4,033 terms
EcoLexicon: Frame-Based Terminology

Terminological Knowledge Base:

- Environmental event is conceptualized as a dynamic process that is initiated by an agent. This affects a patient and produces a result.

- Records: Terms, Graphical Resources and Domains
EcoLexicon

Water

Stream

Capillary water
Termontography

• Multidisciplinary and Functional Approach:
  – Terminography
  – Ontology engineering

• Knowledge-rich Contexts retrieved from corpora
Knowledge-rich Context

“Knowledge-rich contexts are naturally occurring utterances that explicitly describe attributes of domain-specific concepts or semantic relations holding between them at a certain point in time, in a manner that is likely to help the reader of the context understand the concept in question.” (Schumann, 2012)
Knowledge-rich Context

• Knowledge-rich contexts are extracted by knowledge base management systems that are used for creating and displaying relationships between units of information.

• Methods: lexical language patterns; contexts; semantic relations; knowledge patterns; knowledge probes; bag-of-equivalents; term alignment approach
Example: FF POIROT

FF POIROT

Ontology and Fraud

Tools
Critical Evaluation

Terminology Proper → Documentation

Terminography → Linguistics deals with Language and not with Information Tools

Knowledge Engineering → Prototypes
Critical Evaluation

• Theory must focus on the nature of information tools:
  – User’s needs Always punctual needs in specific usage situations
  – Access routes Always quick and easy routes
  – Data Always prepared for allowing users to convert data into Information
Summary

• Terminology has been ‘kidnapped’ by linguists and knowledge engineers.

• Linguists and knowledge engineers are wrong in their assumptions.

• Terminological/knowledge bases do not deliver what they claim.